FREQUENCY BASED MODULATOR COMPENSATION

BACKGROUND OF THE INVENTION

[0001]   Power amplifiers (PAs) are indispensable components in communication systems and are nonlinear in nature. At the PA output, the nonlinearity not only causes spectral regrowth, which interferes with adjacent channel signals, but also degrades the bit error rate of the inband signal. To compensate for the nonlinearity, PA linearization is often necessary.

[0002]   Among the linearization techniques, digital baseband predistortion is highly cost effective. It adds a functional block, called a predistorter, before the digital-to-analog conversion and RF upconversion but after any other baseband processing. The predistorter ideally applies the exact inverse response of the power amplifier to a scaled version of the input signal. Therefore, the output of the predistorter-PA cascade is the input signal multiplied by a scaling factor. To construct such a predistorter, a feedback path is needed to capture the output of the PA.

[0003]   The performance of baseband predistortion relies on accurate PA modeling and obtaining a precise inverse of the PA. In reality, however, the performance can also be affected significantly by imperfections in the upconverter in the transmitter and the downconverter in the feedback path. These imperfections are caused by the analog components employed in both the upconverter and the downconverter, such as mixers, filters, quadrature modulators, and quadrature demodulators.

[0004] To reduce these impairments in the downconverter, filters that have relatively flat frequency may be used, such as LC bandpass and lowpass filters, and digital demodulation, which is free of any demodulation errors. With these configurations and careful design, the imperfections in the downconverter can be negligible.

[0005] In the upconverter, there are two common configurations, but the analog impairments normally cannot be neglected in either. The first uses digital modulation and two-stage upconversion, i.e., first upconvert baseband signals to IF and then to RF. Because of the stringent image rejection requirements of the transmitter, a SAW filter is usually used in the IF stage for this configuration. But the SAW filter often has large frequency response variations, and therefore distorts the predistorted signal. The second choice for the upconverter is to use direct upconversion, in which the I/Q data streams are directly modulated to RF. This structure enables the upconverter to be easily reconfigured to generate RF signals in different frequency bands. It also uses fewer components and is easier to integrate. However, in practice, the quadrature carriers in the analog modulator do not have exactly the same amplitudes and an exact phase difference of 90 degrees. These effects are called gain/phase imbalance and cause cross-talk between the I and Q channels. In addition, leakage of the carriers to the transmitted signal manifests itself in the demodulated received signal as a dc offset. If uncompensated, the gain/phase imbalance and the dc offset would have a large effect on predistortion performance. Various techniques have been proposed in the past to compensate for these impairments. In

these techniques, the gain/phase imbalance are assumed to be frequency independent. However, experiments performed by the inventors indicate that the gain/phase imbalance exhibits frequency dependent behavior when the signal bandwidth becomes wide, for example, 15 MHz.

SUMMARY OF THE INVENTION

[0006]   The present invention provides a transmitter that includes an upconverter for converting a lower frequency signal to a higher frequency signal and a compensator that compensates for at least gain/phase distortion introduced into the lower frequency signal by at least the upconversion. The compensator may also compensate for dc offset introduced into the lower frequency signal.

[0007]   In one exemplary embodiment, the upconverter is a direct upconverter that directly upconverts a baseband signal to the RF signal, and the baseband signal includes in-phase and quadrature phase components. The compensator includes filters compensating for gain/phase imbalance in the in-phase components and gain/phase imbalance in the quadrature phase components.

[0008]   In another exemplary, the transmitter includes a compensator constructor that, based on a model of at least the direct upconverter including an in-phase channel, a quadrature phase channel, and cross coupling channels between the in-phase and quadrature phase channels, estimates the in-phase channel, the quadrature phase channel, and the cross coupling channels between the in-phase and quadrature phase

channels. The compensator constructor constructs the filters in the compensator based on the estimated channels.

[0009]   In another embodiment, at least one filter to correct for gain/phase distortion introduced by an upconverter is derived based on an inverse of the channel model for an upconverter. The inverse of the channel model for the upconverter is based on a cost function, which represents a mean square error, in the frequency domain, between a desired response of a system including at least the upconverter and an actual response of the system including at least the filter and the upconverter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010]   The present invention will become more fully understood from the detailed description given herein below and the accompanying drawings which are given by way of illustration only, wherein like reference numerals designate corresponding parts in the various drawings, and wherein:

[0011]   Fig. 1 illustrates the general structure of transmitter with a baseband predistortion system according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE EXEMPLARY EMBODIMENTS

Channel Models

[0012]   Fig. 1 illustrates the general structure of transmitter with a baseband predistortion system according to an embodiment of the present invention. As shown, a main signal path includes a predistorter 10, an I/Q

compensator 12, a digital-to-analog converter (DAC) 14, a direct upconverter 16 and a power amplifier 18 connected in series. A coupler 20 couples the signal output from the power amplifier 18 to form a feedback path that includes a down converter 22, an analog-to-digital converter (ADC) 24, a digital demodulator 26 and a predistorter design unit 28. As shown in phantom lines, a secondary coupler 30 selectively delivers the output of the direct upconverter 16 to the down converter 22, and an I/Q compensator constructor 32 selectively configures the taps of the I/Q compensator 12. Except for the I/Q compensator 12, secondary coupler 30 and I/Q compensator construction unit 32, the other elements of the baseband predistortion system are well-known in the art. For the purposes of brevity, these elements will not be described in detail.

[0013] As discussed in the background section, the down converter 22 has a relatively flat frequency response and the digital demodulator 26 is free of demodulation errors. Accordingly, the feedback path causes no extra distortion except for additive white noise. In the main path (i.e., the transmitter path), direct upconversion has been adopted, which may be readily reconfigured to generate RF signals in different frequency bands, but introduces gain/phase imbalance and dc offset. The additional baseband processing for the input signal u(n) in the transmitter includes predistortion and I/Q compensation, whose outputs are denoted by z(n) and x(n), respectively.

[0014] During an initialization phase, the predistorter 10 and the I/Q compensator 12 are bypassed (i.e., x(n) = z(n) = u(n)) by having the I/Q

compensator constructor 32 output a training signal to the DAC 14. The secondary coupler 30 sends the output of the direct upconverter 16 to the down converter 22, bypassing the PA 18, during this initialization phase. Based on x(n) and the output y(n) from the digital demodulator 26, the I/Q compensation constructor 32 estimates the parameters of a channel model as discussed in detail below and determines the taps of the filters forming the I/Q compensator 12 as discussed in detail below. After the initialization phase, the bypassed elements are inserted and the predistorter 10 is trained in the well-known manner.

[0015] As will be appreciated the initialization may be performed in many ways. For example, the predistorter 10 may be kept in the loop. Here, the secondary coupler 30 is not needed, as the output of the power amplifier 18 is received by the down converter 22 via the coupler 20. Based on the input to the predistorter 10 and the output y(n) from the digital demodulator 26, the I/Q compensation constructor 32 estimates the parameters of a channel model as discussed in detail below and determines the taps of the filters forming the I/Q compensator 12 as discussed in detail below.

[0016] During operation, the I/Q compensator constructor 32 may periodically or upon command re-estimate the received parameters using the compensation signal x(n) and the output of the digital demodulator y(n) and determine new taps of the filters forming the I/Q compensator 12. In this manner, the I/Q compensator 12 dynamically adapts to changes in the gain/phase imbalance and/or dc offset.

[0017] Next the method of modeling the channel creating the gain/phase imbalance and dc offset will be described in detail, followed by a description of estimating the parameters of the channel model. Lastly, determining the taps of the filters forming the I/Q compensator will be described. As part of the above-described aspects of the present invention, a detailed structure of the I/Q compensator according to one embodiment of the present invention will be provided.

[0018] A detailed view of the channel from $x(n)$ to $y(n)$, is shown in Fig. 2, where Re{ ·} and Im{ ·} denote the real and imaginary parts of a complex number, respectively. We also use subscripts i (in-phase) and q (quadrature) to denote the real and imaginary parts of a complex sequence; for example, we have

$$x(n) = x_i(n) + jx_q(n), \quad y(n) = y_i(n) + jy_q(n).$$

[0019] Note that there may be a fixed delay and phase rotation (caused by the phase difference between the LOs) between $x(n)$ and $y(n)$, but they can be easily removed in any well-known manner.

Real I/Q Channel Model

[0020] The frequency dependent gain/phase imbalance affects the I and Q channels and the cross coupling between them. Moreover, the analog lowpass filters (LPF) and the mixers on the I and Q paths in Fig. 2 may not be exactly the same. To model the I and Q channels and the cross coupling channels between them, we use four real filters, h_{11}, h_{12}, h_{21}, and h_{22},, where h_{11} and h_{22} represent the I and Q channels, h_{12} represents the cross

coupling of the Q channel with the I channel, and h_{21} represents the cross coupling of the I channel with the Q channel. The channel output $y(n)$ then can be written as

$$y(n) = \sum_{k=0}^{K-1} \left\{ \left[x_i(n-k)h_{11}(k) + x_q(n-k)h_{12}(k) \right] + j \left[x_q(n-k)h_{22}(k) + x_i(n-k)h_{21}(k) \right] \right\} + d_i + w_i(n) + j \left[d_q + w_q(n) \right] \tag{1}$$

where $w_i(n)$ and $w_q(n)$ are, respectively, the real and imaginary parts of the complex white noise $w(n)$ and d_i and d_q are, respectively, the real and imaginary parts of the dc offset d. In (1), we assume that all four filters have the same length K. This assumption simplifies the model and the derivations in the following sections although the same methodology can still be applied if the h filters have different lengths.

Complex I/Q Channel Model

[0021]   In (1), we can combine the terms that have $x_i(n\text{-}k)$ or $x_q(n\text{-}k)$ and rewrite in a more compact form; i.e.,

$$y(n) = \sum_{k=0}^{K-1} \left[x_i(n-k)h_i(k) + x_q(n-k)h_q(k) \right] + d + w(n), \tag{2}$$

where

$$h_i(k) = h_{11}(k) + jh_{21}(k), \quad h_q(k) = h_{12}(k) + jh_{22}(k).$$

[0022]   In other words,

$$h_{11}(k) = \text{Re}\{h_i(k)\}, \quad h_{12}(k) = \text{Re}\{h_q(k)\},$$

$$h_{21}(k) = \text{Im}\{h_i(k)\}, \quad h_{22}(k) = \text{Im}\{h_q(k)\}.$$

[0023]   Note that if there is no dc offset in the system and $h_q(k) = jh_i(k)$, $y(n)$ can also be written as

8

$$y(n) = \sum_{k=0}^{K-1} x(n-k)h(k) + w(n), \qquad (3)$$

where $h(k) = h_i(k) = -jh_q(k)$. In this special case, the model in (2) is simplified to a single complex filter with coefficients $h(k)$. This is the case when a digital IF (intermediate frequency) modulator (a form of upconverter that converts a baseband signal to an intermediate frequency signal) is used instead of a direct conversion modulator or an analog IF modulator. Here, the error introduced by the IF modulator and subsequent RF upconverter for converting the IF signal to an RF signal is more broadly thought of as gain/phase distortion, which in the case where $h(k) = h_i(k) = -jh_q(k)$ does not hold true includes gain/phase imbalance.

Direct/Image Channel Model

[0024] The complex I/Q model in (2) gives the relationship between the system output $y(n)$ and the real and imaginary parts of the system input $x(n)$. However, it is not clear from (2) how the channel affects the input $x(n)$ as a whole. We know that

$$x_i(n) = \frac{x(n) + x^*(n)}{2}, \ x_q(n) = \frac{x(n) - x^*(n)}{2j}, \qquad (4)$$

where $(\cdot)^*$ denotes complex conjugate. Substituting (4) in (2), we obtain

$$y(n) = \sum_{k=0}^{K-1} \left[\frac{x(n-k) + x^*(n-k)}{2} h_i(k) + \frac{x(n-k) - x^*(n-k)}{2j} h_q(k) \right] + d + w(n). \quad (5)$$

Rearrange the r.h.s. of (5) to write

$$y(n) = \sum_{k=0}^{K-1} \left[x(n-k)h_d(k) + x*(n-k)h_m(k) \right] + d + w(n) \,,$$

(6)

where

$$h_d(k) = \frac{h_i(k) - jh_q(k)}{2}, \quad h_m(k) = \frac{h_i(k) + jh_q(k)}{2} \,,$$
(7)

are, respectively, the *direct* and *image* transfer function. Next we look at the relationship between *x(n)* and *x*(n)*. If *x(n)* is a lower sideband signal, it can be expressed as the summation of a real sequence *x_d(n)* and its Hilbert transform $\hat{x}_d(n)$, i.e.,

$$x(n) = x_d(n) + j\hat{x}_d(n) \,.$$
(8)

[0025]   Since *x_d(n)* is a real sequence, it has a double sideband frequency spectrum $X_d(e^{j\omega})$. The spectrum of *x(n)* is actually the lower sideband of $X_d(e^{j\omega})$ but with twice the amplitude. From (8), it follows that $x*(n) = x_d(n) - j\hat{x}_d(n)$, whose spectrum is the upper sideband of $X_d(e^{j\omega})$ with twice the amplitude. Similarly, when *x(n)* has an upper sideband spectrum, the spectrum of *x*(n)* is located at the lower sideband. For this reason, we call *x*(n)* the image of *x(n)*, *h_m(k)* the image channel transfer function, and *h_d(k)* the direct channel transfer function. In the channel model (6), we can see that if *h_i(k) ≠ h_q(k)*, i.e., *h_m(k) ≠ 0*, and the input *x(n)* has one-sided spectrum, the image of *x(n)* will show up in the other side of the spectrum, and its amplitude is determined by *h_m(k)*. When the input *x(n)* has a double-

sided spectrum, the images of lower and upper sidebands of $X(e^{j\omega})$ will also appear in the spectrum of the output; however, the images may be completely covered up by the spectrum of the original input signal in that case.

[0026] Note that the three models proposed are all equivalent models. However, they reveal different aspects of the channel, and each has their own pros and cons.

Least-Squares Method of Channel Estimation

[0027] For a block of N $x(n)$ and $y(n)$ data samples, (2) can be written in vector form; i.e.,

$$y = X_i h_i + X_q h_q + d1_P + w, \qquad (9)$$

where $y = \begin{bmatrix} y(K-L-1) & y(K-L)...y(N-L-1) \end{bmatrix}^T$ is a vector length

$$P = N - K + 1 \qquad (10)$$

with L a selectable delay, $X_i = \mathrm{Re}(X)$ and $X_q = \mathrm{Im}(X)$ with

$$X = \begin{bmatrix} x(K-1) & x(K-2) & x(K-3) & \cdots & x(0) \\ x(K) & x(K-1) & x(K-2) & \cdots & x(1) \\ \vdots & \vdots & \vdots & & \vdots \\ x(N-1) & x(N-2) & x(N-3) & \cdots & x(N-K) \end{bmatrix}, \qquad (11)$$

$h_i = [h_i(0)h_i(1)\cdots h_i(k-1)]^T$, $h_q = \begin{bmatrix} h_q(0) & h_q(1)\cdots h_q(K-1) \end{bmatrix}^T$, 1_P is a column vector filled with all ones, and $w = [\omega(K-1)\ \omega(K)\ ...\ w(N-1)]^T$. Since K is the length of the model channel filter, we choose

$$L = \lfloor (K-1)/2 \rfloor, \qquad (12)$$

where $\lfloor x \rfloor$ is the largest integer that is less than or equal to x, such that the maximum response in the model channel filter is located approximately at the center tap. Note that not all N samples of $y(n)$ are used in the formulation in order to avoid the boundary effect. The system output $y(n)$ is nominally matched with the input $x(n)$, i.e., the relative delay, amplitude, and phase difference between $y(n)$ and $x(n)$ have been removed.

[0028] To estimate the channel coefficients, we define a cost function as the mean square difference between the actual output y and the noiseless model output:

$$J = (y - X_i h_i - X_q h_q - d1_P)^H (y - X_i h_i - X_q h_q - d1_P), \qquad (13)$$

where $(\cdot)^H$ denotes Hermitian transpose. The optimal h_i, h_q, and d that minimize the cost function can be found by setting the partial derivatives of J with respect to h_j^*, h_q^*, and d^* to zero (assuming that h_i, h_q and d are constants; i.e.,

$$\frac{\partial J}{\partial h_i^*} = X_i^H (y - X_i h_i - X_q h_q - d1_P) = 0 \qquad (14)$$

$$\frac{\partial J}{\partial h_q^*} = X_q^H (y - X_i h_i - X_q h_q - d1_P) = 0 \qquad (15)$$

$$\frac{\partial J}{\partial d^*} = 1_P^T (y - X_i h_i - X_q h_q - d1_P) = 0 \qquad (16)$$

[0029] Since Xi and Xq in (14) and (15) are real matrices, the Hermitian transpose can be replaced with a transpose without conjugate. Rearranging (14) - (16) and putting them in a single equation, we have

$$\begin{bmatrix} X_i^T X_i & X_i^T X_q & X_i^T 1_P \\ X_q^T X_i & X_q^T X_q & X_q^T 1_P \\ 1_P^T X_i & 1_P^T X_q & 1_P^T 1_P \end{bmatrix} \begin{bmatrix} h_i \\ h_q \\ d \end{bmatrix} = \begin{bmatrix} X_i^T y \\ X_q^T y \\ 1_P^T y \end{bmatrix}. \tag{17}$$

[0030] Therefore, the least-squares estimates of the h_i, h_q, and d are

$$\begin{bmatrix} \hat{h}_i \\ \hat{h}_q \\ \hat{d} \end{bmatrix} = \begin{bmatrix} X_i^T X_i & X_i^T X_q & X_i^T 1_P \\ X_q^T X_i & X_q^T X_q & X_q^T 1_P \\ 1_P^T X_i & 1_P^T X_q & 1_P^T 1_P \end{bmatrix}^{-1} \begin{bmatrix} X_i^T y \\ X_q^T y \\ 1_P^T y \end{bmatrix} \tag{18}$$

[0031] Note that a coarse estimate of the dc offset can be obtained as the difference between the mean of the system output and that of the system input; i.e.,

$$\hat{d} = \frac{1}{N} \sum_{n=0}^{N-1} [y(n) - x(n)]. \tag{19}$$

where N is the window of samples over which the dc offset is determined. From the complex I/Q model, we know that d is given by

$$d = d_y - \left[\mathrm{Re}\{d_x\} \sum_{k=0}^{K-1} h_i(k) + \mathrm{Im}\{d_x\} \sum_{k=0}^{K-1} h_q(k) \right], \tag{20}$$

where d_x and d_y are, respectively, the mean values of the input $x(n)$ and the output $y(n)$. Therefore, in order for (19) to be accurate, either d_x needs to be very small or $\sum_{k=0}^{K-1} h_i(k) \approx 1$ and $\sum_{k=0}^{K-1} h_q(k) \approx j$. If either of these conditions

13

hold, then we may use (19) to estimate the dc offset and adjust the cost function (13) to

$$J = (y - \hat{d}1_P - X_i h_i - X_q h_q)^H (y - \hat{d}1_P - X_i h_i - X_q h_q), \qquad (21)$$

[0032] Similar to the previous derivation of the least-squares solutions, the least-squares estimates of the h_i, h_q are then given by

$$\begin{bmatrix} \hat{h}_i \\ \hat{h}_q \end{bmatrix} = \begin{bmatrix} X_i^T X_i & X_i^T X_q \\ X_q^T X_i & X_q^T X_q \end{bmatrix}^{-1} \begin{bmatrix} X_i^T (y - \hat{d}1_P) \\ X_q^T (y - \hat{d}1_P) \end{bmatrix} \qquad (22)$$

Least-Squares Method with Diagonal Loading

[0033] Since the input signal $x(n)$ is usually a bandpass signal, the channel estimates are only accurate within the band. The out-of-band responses are solely determined by the wideband noise in $y(n)$. These arbitrary out-of-band responses are not desired in predistortion applications, where the low level out-of-band signal needs to be accurately preserved in order to compensate for the PA nonlinearity. A simple solution is to add an artificial white noise to both $x(n)$ and $y(n)$ to create a flat frequency response over the whole band. The artificial noise level may be controlled to be well above the original white noise level in $y(n)$ but well below the inband signal level. Thus, an artificial flat out-of-band response is created that preserves the inband channel response. Let us denote the artificial white noise as $v(n)$, whose real and imaginary parts are, respectively, $v_i(n)$ and $v_q(n)$. Similar to the definition of X_i, and X_q, we can

14

define V_i and V_q and rewrite (18) for the case where $v(n)$ is added to both $x(n)$ and $y(n)$, which leads to

$$
\begin{bmatrix} \hat{h}_i \\ \hat{h}_q \\ \hat{d} \end{bmatrix} = \begin{bmatrix} (X_i+V_i)^T(X_i+V_i) & (X_i+V_i)^T(X_q+V_q) & (X_i+V_i)^T 1_P \\ (X_q+V_q)^T(X_i+V_i) & (X_q+V_q)^T(X_q+V_q) & (X_q+V_q)^T 1_P \\ 1_P^T(X_i+V_i) & 1_P^T(X_q+V_q) & 1_P^T 1_P \end{bmatrix}^{-1}
$$

$$
\times \begin{bmatrix} (X_i+V_i)^T(y+v) \\ (X_q+V_q)^T(y+v) \\ 1_P^T(y+v) \end{bmatrix}, \tag{23}
$$

where $v = [v(K-1)\ v(K)\ \ldots\ v(N-1)]^T$. Ideally, $v_i(n)$ and $v_q(n)$ are uncorrelated with each other and are uncorrelated with $x(n)$ and $y(n)$. Moreover, the mean value of $v(n)$ is zero. Therefore, instead of calculating (23), we may estimate the channel parameters using the following formula

$$
\begin{bmatrix} \hat{h}_i \\ \hat{h}_q \\ \hat{d} \end{bmatrix} = \begin{bmatrix} X_i^T X_i + \sigma^2 I & X_i^T X_q & X_i^T 1_P \\ X_q^T X_i & X_q^T X_q + \sigma^2 I & X_q^T 1_P \\ 1_P^T X_i & 1_P^T X_q & 1_P^T 1_P \end{bmatrix}^{-1} \begin{bmatrix} X_i^T y + \sigma^2 e \\ X_q^T y + \sigma^2 e \\ 1_P^T y \end{bmatrix}, \tag{24}
$$

where σ^2 is the variance of the artificial noise, I is a $K \times K$ identity matrix, and $e = \begin{bmatrix} 0_L^T & 1 & 0_L^T \end{bmatrix}^T$, where 0_L and 0_M are, respectively, length $L = \lfloor (K-1)/2 \rfloor$ and $M = K - L - 1$ column vectors filled with all zeros. In the case where the dc offset is estimated using (19), the least-squares estimates of the h_i, h_q with diagonal loading reduce to

$$
\begin{bmatrix} \hat{h}_i \\ \hat{h}_q \end{bmatrix} = \begin{bmatrix} X_i^T X_i + \sigma^2 1 & X_i^T X_q \\ X_q^T X_i & X_q^T X_q + \sigma^2 I \end{bmatrix}^{-1} \begin{bmatrix} X_i^T(y-\hat{d}1_P) + \sigma^2 e \\ X_q^T(y-\hat{d}1_P) + \sigma^2 e \end{bmatrix} \tag{25}
$$

[0034] In either case, the diagonal loading has the additional advantage of regularizing the solution, i.e., reducing the condition number of the correlation matrix, so that more accurate solutions can be achieved.

[0035] As will be appreciated from the above description, the I/Q compensator constructor 32 estimates the dc offset from the signals $x(n)$ and $y(n)$ using equation (19) and estimates the response of the I or in-phase channel $\hat{h}_i$ and the response of the Q or quadrature phase channel $\hat{h}_q$ from the signals $x(n)$ and $y(n)$ using either equation (22) or equation (25).

Construction of the I/Q Compensator

[0036] After obtaining the channel filter estimates $\hat{h}_i$ and $\hat{h}_q$ and the dc offset estimate $\hat{d}$, the I/Q compensator 12 is constructed by the I/Q compensator constructor 32 to compensate for these imperfections. It turns out that these imperfections can be fully compensated by an I/Q compensator that has structure similar to the channel models described above. To ease the derivation, one exemplary embodiment uses the real I/Q channel model to represent both the I/Q compensator 12 and the baseband direct upconverter 16. A block diagram of the cascade of the I/Q compensator 12 and the upconverter 16 is shown in Fig. 3. As shown, the I/Q compensator 12 includes four real finite impulse response (FIR) filters, g_{11}, g_{12}, g_{21}, g_{22}. The first filter g_{11} filters the in-phase components and the second filter g_{12} filters the quadrature phase components. A first adder 50 adds the outputs of the first and second filters g_{11} and g_{12} to produce a gain/phase compensated in-phase signal. The third filter g_{21} filters the in-

phase components and the fourth filter g_{22} filters the quadrature phase components. A second adder 52 adds the outputs of the third and fourth filters g_{21} and g_{22} to produce a gain/phase compensated quadrature phase signal. A third adder 54 adds an in-phase dc component c_i to the in-phase signal and a fourth adder 56 adds a quadrature phase dc component c_q to the quadrature phase signal to compensate for dc offset.

Offset Compensation

[0037] Let us first consider the compensation of the dc offset $\hat{d}$ using the dc component (c_i, c_q) in the I/Q compensator 12. After passing through the four channel filters g_{11}, g_{12}, g_{21}, and g_{22}, the dc component c becomes

$$
\begin{aligned}
\tilde{c}_i &= c_i s_{11} + c_q s_{12} \\
\tilde{c}_q &= c_i s_{21} + c_q s_{22}
\end{aligned}
\tag{26}
$$

where

$$
\begin{aligned}
s_{11} &= \sum_{k=0}^{K-1} \hat{h}_{11}(k), \quad s_{12} = \sum_{k=0}^{K-1} \hat{h}_{12}(k) \\
s_{21} &= \sum_{k=0}^{K-1} \hat{h}_{21}(k), \quad s_{22} = \sum_{k=0}^{K-1} \hat{h}_{22}(k)
\end{aligned}
\tag{27}
$$

[0038] Thus, adding c in the I/Q compensation 12 is equivalent of adding a dc component $\tilde{c}$ at the end of the cascade. In order to compensate for the dc offset $\tilde{d}$, we would like to have $\tilde{c}_i = -\hat{d}_i$ and $\tilde{c}_q = -\hat{d}_q$; i.e.,

$$
\begin{aligned}
c_i s_{11} + c_q s_{12} &= -\hat{d}_i \\
c_i s_{21} + c_q s_{12} &= -\hat{d}_q
\end{aligned}
\tag{28}
$$

It is clear from (28) that the solution for c_i and c_q is

$$\begin{bmatrix} c_i \\ c_q \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}^{-1} \begin{bmatrix} -\hat{d}_i \\ -\hat{d}_q \end{bmatrix}. \qquad (29)$$

[0039]   Note that when the inverse matrix in (29) is close to the identity matrix, we have

$$\begin{bmatrix} c_i \\ c_q \end{bmatrix} \approx \begin{bmatrix} -\hat{d}_i \\ -\hat{d}_q \end{bmatrix}. \qquad (30)$$

[0040]   Accordingly, the I/Q compensator constructor 32 sets the dc offset compensation c_i, c_q according to equation (29) or (30).

Gain/Phase Imbalance Compensation

[0041]   Next we consider the problem of designing the four filters g_{11}, g_{12}, g_{21}, and g_{22} in the I/Q compensator 12. We adopt a frequency domain approach to solve this problem. Denote the Fourier transform of $y_i(n)$ and $y_q(n)$ by $Y_i(e^{j\omega})$ and $Y_q(e^{j\omega})$, respectively, and the Fourier transform of $z_i(n)$ and $z_q(n)$ by $Z_i(e^{j\omega})$ and $Z_q(e^{j\omega})$, respectively. The outputs $Y_i(e^{j\omega})$ and $Y_q(e^{j\omega})$ of the digital demodulator 26 can be expressed as

$$\begin{bmatrix} Y_i(e^{j\omega}) \\ Y_q(e^{j\omega}) \end{bmatrix} = \begin{bmatrix} \hat{H}_{11}(e^{j\omega}) & \hat{H}_{12}(e^{j\omega}) \\ \hat{H}_{21}(e^{j\omega}) & \hat{H}_{22}(e^{j\omega}) \end{bmatrix} \begin{bmatrix} G_{11}(e^{j\omega}) & G_{12}(e^{j\omega}) \\ G_{21}(e^{j\omega}) & G_{22}(e^{j\omega}) \end{bmatrix} \begin{bmatrix} Z_i(e^{j\omega}) \\ Z_q(e^{j\omega}) \end{bmatrix}. \qquad (31)$$

To make $Y_i(e^{j\omega}) = Z_i(e^{j\omega})$ and $Y_q(e^{j\omega}) = Z_q(e^{j\omega})$, we need

$$\begin{bmatrix} G_{11}(e^{j\omega}) & G_{12}(e^{j\omega}) \\ G_{21}(e^{j\omega}) & \hat{G}_{22}(e^{j\omega}) \end{bmatrix} = \begin{bmatrix} \hat{H}_{11}(e^{j\omega}) & \hat{H}_{12}(e^{j\omega}) \\ \hat{H}_{21}(e^{j\omega}) & \hat{H}_{22}(e^{j\omega}) \end{bmatrix}^{-1}$$

$$= A(e^{j\omega}) \begin{bmatrix} \hat{H}_{22}(e^{j\omega}) & -\hat{H}_{12}(e^{j\omega}) \\ -\hat{H}_{21}(e^{j\omega}) & \hat{H}_{11}(e^{j\omega}) \end{bmatrix}, \qquad (32)$$

where

$$A(e^{j\omega}) = \frac{1}{\left[\hat{H}_{11}(e^{j\omega})\hat{H}_{22}(e^{j\omega}) - \hat{H}_{12}(e^{j\omega})\hat{H}_{21}(e^{j\omega}) \right]}. \qquad (33)$$

[0042] In the time domain, we may view (32) as follows,

$$\begin{aligned} g_{11}(k) &= \hat{h}_{22}(k)*a(k), \quad g_{12}(k) = -\hat{h}_{12}(k)*a(k) \\ g_{21}(k) &= -\hat{h}_{21}(k)*a(k), \quad g_{22}(k) = \hat{h}_{11}(k)*a(k), \end{aligned} \qquad (34)$$

where * denotes convolution and $a(k)$ is the inverse Fourier transform of $A(e^{j\omega})$. From the definition of $A(e^{j\omega})$, $a(k)$ is the inverse of filter

$$h_c(k) = \hat{h}_{11}(k)*\hat{h}_{22}(k) - \hat{h}_{12}(k)*\hat{h}_{21}(k) ; \qquad (35)$$

i.e., the convolved response of filters $a(k)$ and $h_c(k)$, denoted by $c(k)$, approximates a delta function with a certain delay. In other words, we would like the Fourier transform of $c(k)$ to be close to a desired frequency response $e^{-j\omega n_0}$, where n_0 is the desired delay. To find the optimal $a(k)$ coefficients, we suppose that filter $a(k)$ has K_a taps, the convolved response of $a(k)$ and $h_c(k)$ can be written as

$$c(k) = \sum_{l=0}^{K_a-1} h_c(k-l)a(l), k = 0, 1, \cdots, K_a + 2K - 3, \qquad (36)$$

which has length $K_c = K_a + 2K - 2$. [Note that filter $h_c(k)$ is of length $2K-1$]. We then define a cost function as the integrated difference between the

frequency response of filter *c(k)* and the desired frequency response within given passband $[-\omega_P, \omega_P]$; i.e.,

$$J_c = \int_{-\omega_P}^{\omega_P} \left[\sum_{k_1=0}^{K_c-1} c(k_1) e^{-j\omega k_1} - e^{-j\omega n_0} \right] \left[\sum_{k_2=0}^{K_c-1} c(k_2) e^{-j\omega k_2} - e^{-j\omega n_0} \right]^* d\omega. \quad (37)$$

[0043] Substituting (36) into (37), we have

$$J_c = \int_{-\omega_P}^{\omega_P} \left[\sum_{k_1=0}^{K_c-1} \sum_{l_1=0}^{K_a-1} h_c(k_1 - l_1) a(l_1) e^{-j\omega k_1} - e^{-j\omega n_0} \right]$$

$$\times \left[\sum_{k_2=0}^{K_c-1} \sum_{l_2=0}^{K_a-1} h_c(k_2 - l_2) a(l_2) e^{-j\omega k_2} - e^{-j\omega n_0} \right]^* d\omega. \quad (38)$$

[0044] To find the optimal *a(k)* that minimizes the cost function, we take the partial derivative of J_c in (38) with respect to *a*(l)* (assuming *a(l)* is constant and set it to zero); i.e.,

$$\frac{\partial J_c}{\partial a^*(l)} = \int_{-\omega_P}^{\omega_P} \left[\sum_{k_1=0}^{K_c-1} \sum_{l_1=0}^{K_a-1} h_c(k_1 - l_1) a(l_1) e^{-j\omega k_1} - e^{-j\omega n_0} \right]$$

$$\times \sum_{k_2=0}^{K_c-1} h_c^*(k_2 - 1) e^{j\omega k_2} d\omega = 0; \quad l = 0, \ldots, K_a - 1. \quad (39)$$

[0045] Rearrange (39) to write

$$\frac{\partial J_c}{\partial a^*(l)} = \sum_{l_1=0}^{K_a-1} a(l_1) \sum_{k_1=0}^{K_c-1} \sum_{k_2=0}^{K_c-1} h_c(k_1 - l_1) h_c^*(k_2 - l) \int_{-\omega_P}^{\omega_P} e^{-j\omega(k_1 - k_2)} d\omega$$

$$- \sum_{k_2=0}^{K_c-1} h_c^*(k_2 - l) \int_{-\omega_P}^{\omega_P} e^{-j\omega(n_0 - k_2)} d\omega = 0. \quad (40)$$

[0046] In predistortion applications, the least-squares with diagonal loading is used for estimating the four *h* filters, and filters $h_{11}(k)$ and $h_{22}(k)$ have flat responses outside of the signal band. Moreover, the responses of

the cross coupling filters $h_{12}(k)$ are usually much smaller than those of the filters $h_{11}(k)$ and $h_{22}(k)$. Therefore, the out-of-band response of filter $h_c(k)$ may be considered as flat with unit gain. Thus, we can use $\omega_P = \pi$ in (40). In the channel estimation phase, we want the maximum amplitudes of the filters to occur around the center taps. When we construct the inverse filters here, we also let our desired response for filter $c(n)$ be a delta function around the center tap; i.e., $n_0 = (K_c - 1)/2$. Substituting $\omega_P = \pi$ $n_0 = (K_c - 1)/2$ into (40) and carrying out the integration, we obtain

$$\frac{\partial J_c}{\partial a^*(l)} = \sum_{l_1=0}^{K_a-1} a(l_1) \sum_{k_1=0}^{K_c-1} \sum_{k_2=0}^{K_c-1} h_c(k_1 - l_1) h_c^*(k_2 - l) 2\pi \ \text{sinc}(k_1 - k_2)$$

$$- \sum_{k_2=0}^{K_c-1} h_c^*(k_2 - l) 2\pi \ \text{sinc}(n_0 - k_2) = 0 \,, \tag{41}$$

where

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \tag{42}$$

[0047] For l from 0 to K_a - 1, we have a set of equations from (41); i.e.,

$$\sum_{l_1=0}^{K_a-1} a(l_1) \sum_{k_1=0}^{K_c-1} \sum_{k_2=0}^{K_c-1} 2\pi \, h_c(k_1 - l_1) h_c^*(k_2 - l_2) \ \text{sinc}(k_1 - k_2)$$

$$= \sum_{k_2=0}^{K_c-1} 2\pi \, h_c^*(k_2 - l) \ \text{sinc}(k_1 - k_2). \tag{43}$$

[0048] Removing 2π from both sides of (43) and rewriting it in matrix form, we have

21

$$R_h a = h_a \qquad (44)$$

where a $= \begin{bmatrix} a_0 \dots a_{K_a-1} \end{bmatrix}^T$, and the elements of matrix R_h and column vector h_a are defined as

$$R_h(l_1,l) = \sum_{k_1=0}^{K_c-1}\sum_{k_2=0}^{K_c-1} h_c(k_1 - l_1)h_c^*(k_2 - l)\ \mathrm{sinc}(k_1 - k_2) \qquad (45)$$

$$h_a(l) = \sum_{k_2=0}^{K_c-1} h_c^*(k_2 - l)\ \mathrm{sinc}(n_0 - k_2). \qquad (46)$$

[0049] Note that R_h is a Toeplitz matrix, so we only need to calculate the first row and column of the matrix. The least-squares estimate of **a** from (44) is

$$\hat{a} = R_h^{-1}h_a. \qquad (47)$$

[0050] Accordingly, the I/Q compensator constructor 32 determines $\hat{a}$ according to equation (47) and then determines the filter taps for filters g_{11}, g_{12}, g_{21}, and g_{22}, according to equation (34) using $\hat{a}$ as $a(k)$.

[0051] The output from the power amplifier 18 is then transmitted in a conventional manner. Any well-known receiver for receiving transmitted signals may then receive and demodulate the transmitted signals. For example, the receiver may have the structure of the feedback path shown in Fig. 1.

[0052] In an alternative embodiment, the filter taps for the filters g_{11}, g_{12}, g_{21}, and g_{22} are directly estimated using equation (18) or (24). In using equations (18) and (24), y(n) is treated as the input and x(n) is treated as the

output. The resulting h_i, h_q, and d then form the basis for the I/Q compensator 12.

[0053] The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as a departure from the spirit and scope of the invention, and all such modifications are intended to be included within the scope of the following claims.